Zero-frequency divergence and the gauge phase factor in optical response theory



1999 J. Phys.: Condens. Matter 11 9823

(http://iopscience.iop.org/0953-8984/11/48/338)

J. Phys.: Condens. Matter **11** (1999) 9823–9840. Printed in the UK PII: S0953-8984(99)06050-6

# Zero-frequency divergence and the gauge phase factor in optical response theory

Minzhong Xu† and Xin Sun‡

† Department of Chemistry, New York University, New York, NY 10003, USA
‡ Department of Physics, Fudan University, and National Laboratory of Infrared Physics, Shanghai 200433, People's Republic of China

E-mail: mx200@scivis.nyu.edu (Minzhong Xu)

Received 15 July 1999

**Abstract.** Static current–current correlation leads to a zero-frequency divergence (ZFD) in the definition of optical susceptibilities. Previous computations have shown non-equivalent results for two gauges ($\boldsymbol{p}\cdot\boldsymbol{A}$ and $\boldsymbol{E}\cdot\boldsymbol{r}$) for exactly the same unperturbed wavefunctions. We reveal that these problems are caused by the incorrect treatment of the time-dependent gauge phase factor in optical response theory. The gauge phase factor, which is conventionally ignored by the theory, is important in resolving the ZFD problem and obtaining equivalent results for these two gauges. The Hamiltonians with these two gauges are not necessarily equivalent unless the gauge phase factor is properly considered in the wavefunctions. Both Su–Shrieffer–Heeger (SSH) and Takayama–Lin-Liu–Maki (TLM) models of *trans*-polyacetylene serve as illustrative examples in studying the linear susceptibility $\chi^{(1)}$ through both current–current and dipole–dipole correlations. Previous improper results of $\chi^{(1)}$-calculations and for distribution functions obtained with both gauges are discussed. The importance of the gauge phase factor in solving the ZFD problem is emphasized on the basis of the SSH and TLM models. As a conclusion, the reason for dipole–dipole correlation being preferable to current–current correlation in practical computations is explained.

## 1. Introduction

The static current–current correlation ($J_0 J_0$) [1–3] has been widely applied in optical response theory for many decades, both in the definition of the linear susceptibility $\chi^{(1)}$ or conductivity $\sigma^{(1)}$ [2], and in the definition of the non-linear optical susceptibilities such as $\chi^{(n)}$, where $n \geqslant 2$ [1, 3].

Within the semiclassical theory of radiation, which is also emphasized in the literature, by Mahan and Subbaswamy [2, 4], Butcher and Cotter [3], Bloembergen [5], Shen [6], and Mukamel [7], the electric field is treated classically, and propagation of electromagnetic waves in a medium is governed by Maxwell's equations; that is, the electric field $\boldsymbol{E}(\boldsymbol{r}, t)$ at some specific position $\boldsymbol{r}$ and time $t$ can be described as†

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}_0 \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}-\mathrm{i}\omega t} \tag{1.1}$$

† To guarantee the correctness of the Hermitian of the Hamiltonian, the electrical field $\boldsymbol{E}$ should be expressed as $\boldsymbol{E}_0 \cos(\omega t)$ instead of a complex field [5]. But in the frequency analysis [2–6], it is easier to use the form (1.1) to describe the susceptibilities and this makes virtually no difference to the results. The notation for the definition of the susceptibilities used here follows Butcher and Cotter's notation [3].

0953-8984/99/489823+18$30.00  

where $\boldsymbol{E}_0$ is the amplitude, and $\boldsymbol{k}$ and $\omega$ are the wavevector and frequency. Then, the $\chi^{(n)}$ in unit volume $v$ under the static current–current ($J_0J_0$-) correlation are conventionally defined as follows [1, 3]:

$$\chi^{(n)}(\Omega;\omega_1,\ldots,\omega_n) = -\delta_{n1}\frac{n^{(e)}e^2}{\epsilon_0 m\omega_1^2}\hat{I} + \frac{\chi^{(n)}_{j_0j_0}(\Omega;\omega_1,\ldots,\omega_n)}{\epsilon_0\,\mathrm{i}\Omega\omega_1\cdots\omega_n} \tag{1.2}$$

where

$$\Omega \equiv -\sum_{i=1}^{n}\omega_i.$$

Here, $n^{(e)}$ and $m$ are the electron density and electron mass, $\epsilon_0$ is the dielectric constant, $\hat{I}$ is a unit dyadic, $\delta_{n,1}$ is the Kronecker symbol, and

$$\chi^{(n)}_{j_0j_0}(\Omega;\omega_1,\ldots,\omega_n) = \frac{1}{n!}\left[\frac{1}{\bar{h}}\right]^n\frac{1}{V}\int \mathrm{d}\boldsymbol{r}_1\cdots\mathrm{d}\boldsymbol{r}_n\int \mathrm{d}t_1\cdots\mathrm{d}t_n\int \mathrm{d}\boldsymbol{r}\,\mathrm{d}t\,\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}+\mathrm{i}\Omega t}$$
$$\times\langle\hat{T}\hat{\boldsymbol{J}}_0(\boldsymbol{r},t)\hat{\boldsymbol{J}}_0(\boldsymbol{r}_1,t_1)\cdots\hat{\boldsymbol{J}}_0(\boldsymbol{r}_n,t_n)\rangle \tag{1.3}$$

where $V$ is the total volume, $\hat{T}$ is the time-ordering operator and $\hat{\boldsymbol{J}}_0$ is the static current operator [3].

If we choose the static dipole–dipole ($DD$-) correlation, the $n$th-order susceptibilities will be obtained as follows [3–7]:

$$\chi^{(n)}(\Omega;\omega_1,\ldots,\omega_n) = \frac{1}{n!}\left[\frac{\mathrm{i}}{\bar{h}}\right]^n\frac{1}{V}\int \mathrm{d}\boldsymbol{r}_1\cdots\mathrm{d}\boldsymbol{r}_n\int \mathrm{d}t_1\cdots\mathrm{d}t_n\int \mathrm{d}\boldsymbol{r}\,\mathrm{d}t\,\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}+\mathrm{i}\Omega t}$$
$$\times\langle\hat{T}\hat{\boldsymbol{D}}(\boldsymbol{r},t)\hat{\boldsymbol{D}}(\boldsymbol{r}_1,t_1)\cdots\hat{\boldsymbol{D}}(\boldsymbol{r}_n,t_n)\rangle \tag{1.4}$$

where $\hat{\boldsymbol{D}}$ is the static dipole operator.

It is commonly held that the zero-frequency divergence (ZFD) seen in the static current expression, equation (1.2), is only an apparent problem, and that the gauges $\boldsymbol{E}\cdot\boldsymbol{r}$ and $\boldsymbol{p}\cdot\boldsymbol{A}$ will give exactly the same results for the same unperturbed wavefunctions [8]. In other words, equation (1.2) and equation (1.4) should lead to the same result if one proceeds correctly. In linear response theory, for a homogeneous and isotropic medium, Martin and Schwinger have shown that the ZFD in the conductivity $\sigma^{(1)}$ arising from current expressions can be cancelled by introducing a diamagnetic term (an $\boldsymbol{A}^2$-term) [9]. The cancellation of the ZFD in the linear conductivity by a diamagnetic term is also discussed by Mahan, and Haug and Jauho in their famous books [2, 10], with careful consideration of the limit sequence for $\boldsymbol{k}$ and $\omega$†. In the solid state, for full and empty bands, Aspnes had shown the equivalence of two gauges (equation (1.2) and equation (1.4)) in a demonstration based on the assumption of the cancellation of a ZFD term, equation (2.6), in his seminal work on $\chi^{(2)}$-computations [11]. These works have strengthened the common belief in the equivalence of the two gauges as regards the static dipole and the static current expressions. Also, the notion that the ZFD is merely an apparent problem has usually been accepted with limited justification.

Although the 'apparent' nature of this ZFD has been emphasized by many people, strictly speaking, as far as we are aware, the solution to this ZFD problem is seldom directly obtained from a $J_0J_0$-correlation (like equation (1.2)) unlike in Martin and Schwinger's original proof [9]. As we see in the calculation of $\chi^{(2)}$ for the solid state, the ZFD term is conventionally isolated from the convergent term and is then discarded without a careful direct check [11]. Historically, this ZFD problem has been of no interest in view of the following facts:

† Mahan and Haug suggest a possible sequence limit [2, 10]: firstly setting the wavevector $\boldsymbol{k}\to 0$, and secondly setting the frequency $\omega\to 0$ when considering the linear conductivity $\sigma^{(1)}$ at zero frequency. If the long-wavelength limit applies, the sequence condition for $\boldsymbol{k}$ and $\omega$ is automatically satisfied.

(i) The gauge transformations seem to guarantee the equivalence of the two gauges for the same set of wavefunctions.
(ii) In transport theory [10,12], the correct imaginary part of the $J_0J_0$-correlation (Re$[\sigma^{(1)}(\omega)]$) can still be obtained.
(iii) The ZFD problem (related to Im$[\sigma^{(1)}(\omega)]$) can usually be avoided by applying the Kramers–Kronig (KK) relations to the imaginary part of the $J_0J_0$-correlation [2,10,12,14].

Thus the equivalence of the two gauges has become well accepted in the community, and the ZFD is widely considered to be at most a complex technical problem [2, 10].

Besides the assumptions made in the previous ZFD proofs [2, 3, 9–11], however, in practical applications, there are always some puzzles challenging the above common belief. In a careful study of the choice of gauge for two-photon 1s–2s transitions of the hydrogen atom [8], Bassani, Forney, and Quattropani have found that on directly applying exactly the same unperturbed wavefunctions, the $\boldsymbol{E}\cdot\boldsymbol{r}$ gauge leads to much faster convergence than the $\boldsymbol{p}\cdot\boldsymbol{A}$ gauge when using a limited number of discrete intermediate states; numerical computations also show a 50% difference in transition rate between these gauges if we just include all discrete intermediate states. Thus, Bassani *et al* draw the conclusion that $\boldsymbol{E}\cdot\boldsymbol{r}$ is a better gauge. Also, in the study of oscillator strength in a superlattice by Peeters *et al*, it turns out that $\boldsymbol{E}\cdot\boldsymbol{r}$ and $\boldsymbol{p}\cdot\boldsymbol{A}$ are non-equivalent gauges in the barrier region according to numerical computations [15]. Thus, in contrast to the above, Peeters *et al* concluded that the position operator $\boldsymbol{r}$ in the solid state should be redefined and that $\boldsymbol{p}\cdot\boldsymbol{A}$ is much the better gauge. A recent study on a zinc-blende semiconductor [16] again raises questions regarding the equivalence of the roles of $\boldsymbol{E}\cdot\boldsymbol{r}$ and $\boldsymbol{p}\cdot\boldsymbol{A}$ in representing the transition matrix of the $\chi^{(2)}$-formula†.

Extensive studies on the optical properties of polymers [1,17–20] have been based on tight-binding approximation (TBA) models, such as the Su–Shrieffer–Heeger (SSH) model [21] and the Takayama–Lin-Liu–Maki (TLM) model [22] for weakly correlated systems, and the Hubbard and Pariser–Parr–Pople (PPP) models for strongly correlated systems. These models drastically reduce the complexity of the systems and provide a reasonable way of obtaining real physical insight into many-body systems. For providing gauge invariance, a U(1) phase transformation has been suggested [20, 23] for these models. Because of the use of the static current formula without a diamagnetic term, there will be a ZFD problem in the linear conductivity $\sigma^{(1)}(\omega)$. However, this ZFD problem has in general not been pointed out clearly and it has obviously been neglected in the previous works [17, 20]. To avoid the ZFD in $\sigma^{(1)}$ and to obtain a convergent result, Batistic and Bishop suggest subtracting the term $\langle[j_0, j_0]\rangle(\omega = 0)$ [19], which is supposed to be a diamagnetic term derived directly from the TBA Hamiltonian. Unfortunately, as we will show in this paper, the diamagnetic term derived from the U(1) transformation [20] cannot directly return the expected $\langle[j_0, j_0]\rangle(\omega = 0)$ term. Moreover, the experimentally observed two-photon absorption peak [24,25] in the $\chi^{(3)}$-spectrum of *trans*-polyacetylene has raised wide interest in finding a theoretical explanation. From equation (1.2), on the basis of TLM models, the two-photon cusp has been obtained analytically [1], but this has been criticized by others in view of the dipole formula approach and other physical concerns [26–32]. Recently, a quite different analytical form of the $\chi^{(3)}$-spectrum [33] has been obtained for $DD$-correlation as compared to that for $J_0J_0$-correlation [1]. The above discrepancies have already cast some doubt on the rooted belief of the equivalence of $J_0J_0$- and $DD$-correlations.

In this paper, we will re-examine the concepts of the gauge transformations and directly show that the gauge phase factor, which has not been sufficiently emphasized previously

† In Khurgin and Volsin's work [16], they pointed out the different diagonal matrix elements in the $\chi^{(2)}$-formula that occur for the $\boldsymbol{p}\cdot\boldsymbol{A}$ and $\boldsymbol{E}\cdot\boldsymbol{r}$ gauges.

and is conventionally ignored in the current–current correlation scheme [1, 3], is actually very important in optical response theory for resolving this ZFD difficulty, and recovering equivalent results for the two gauges ($\boldsymbol{p}\cdot\boldsymbol{A}$ and $\boldsymbol{r}\cdot\boldsymbol{E}$). Therefore, the static current operator $\hat{J}_0$ is no longer suitable for considering the equivalence between the two gauges; instead, we should include the induced field currents (IFC) which are introduced by the gauge phase factor. To illustrate the effect of the gauge phase factor, rather than abstract concepts, we choose just the linear response in one-dimensional (1D) periodic TBA models, such as the SSH and TLM models, as our examples. The concepts of the gauge phase factor in the specific linear examples can also be extended to the non-linear optical response theory and two-dimensional (2D) or three-dimensional (3D) cases.

The paper is organized as follows. In section 2, we will re-examine the concept of the gauge transformation and discuss the importance of the gauge phase factor in optical response theory. The problems caused by ignoring the gauge phase factor are discussed within a general scheme, independent of the models. To give an intuitive picture, the linear optical response in periodic TBA models, such as SSH and TLM models, is investigated in this paper. The Hamiltonian for $DD$-correlation ($\boldsymbol{E}\cdot\boldsymbol{r}$ gauge) is discussed (section 3.1), and we will study the linear susceptibility $\chi^{(1)}$ for $DD$-correlation in section 3.2. $\chi^{(1)}$ under current–current correlation is discussed in section 4. The SSH Hamiltonian in the $\boldsymbol{p}\cdot\boldsymbol{A}$ gauge will be obtained in section 4.1. Before applying the gauge phase factor, the ZFD problem of $\chi^{(1)}$ arising from $J_0J_0$-correlation is illustrated in section 4.2; previous qualitatively different solutions for $\chi^{(1)}$ and the practical difficulties in this ZFD problem are also analysed on the basis of the models (section 4.2). After applying the gauge phase factor to the wavefunctions, a convergent result can be obtained and the ZFD problem will be resolved (section 4.3). The conditions for equivalence between the two gauges are discussed in section 5.1 and the influence of the gauge phase factor on the initial distribution function $f_n(\boldsymbol{k})$ for the two gauges is investigated (section 5.2). The reasons for some previous puzzles regarding the choice of gauges are discussed in section 5.3. Conclusions emphasizing the implications of our work will be given in section 6.

## 2. The gauge phase factor in gauge transformation

Gauge transformation is already well understood in optical response theory [34, 35]. The assertion of the equivalence of two gauges rests upon the concept of gauge transformation.

If an electromagnetic field is applied, the Schrödinger equation is given by

$$\mathrm{i}\bar{h}\frac{\partial}{\partial t}\psi(\boldsymbol{r},t)=\left[\frac{1}{2m}(\hat{\boldsymbol{p}}-q\boldsymbol{A})^2+V(\boldsymbol{r})+q\phi\right]\psi(\boldsymbol{r},t) \tag{2.1}$$

where $\psi(\boldsymbol{r},t)$ is the exact wavefunction at space position $\boldsymbol{r}$ and the specific time $t$, $m$ is the particle mass, $q$ is the electrical charge, $V(\boldsymbol{r})$ is the potential, and $\boldsymbol{A}$ and $\phi$ are the vector and scalar potentials respectively, under the following transformation:

$$\begin{aligned}&\boldsymbol{A}\to\boldsymbol{A}'=\boldsymbol{A}+\boldsymbol{\nabla}f(\boldsymbol{r},t)\\&\phi\to\phi'=\phi-\frac{\partial}{\partial t}f(\boldsymbol{r},t)\end{aligned} \tag{2.2}$$

where $f(\boldsymbol{r},t)$ is arbitrary, and $\boldsymbol{A}'$ and $\phi'$ are the new vector and new scalar potentials after the transformation (2.2). It can be shown [35] that the form of the Schrödinger equation will be exactly the same if the old wavefunction $\psi$ undergoes the following change into the new exact wavefunction $\psi'$:

$$\psi\to\psi'=\mathrm{e}^{\mathrm{i}F_g(\boldsymbol{r},t)}\psi=\hat{T}_G(\boldsymbol{r},t)\psi \tag{2.3}$$

where the gauge phase factor $F_g(\boldsymbol{r}, t)$ is defined as

$$F_g(\boldsymbol{r}, t) \equiv \frac{q}{\hbar} f(\boldsymbol{r}, t). \tag{2.4}$$

The above equations (2.2) and (2.3) are called the gauge transformation (or $U(1)$ transformation [23]).

The long-wavelength approximation [2, 3] is used in this paper—that is, we take $\boldsymbol{k} = 0$ in equation (1.1), and the electric field $\boldsymbol{E}$ is described as $\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{-\mathrm{i}\omega t}$.

We consider the following initial scalar and vector potentials in the $\boldsymbol{E} \cdot \boldsymbol{r}$ gauge:

$$\boldsymbol{A} = 0 \qquad \phi = -\boldsymbol{E} \cdot \boldsymbol{r}. \tag{2.5}$$

After choosing the gauge phase factor $F_g$ as

$$F_g = \frac{q\boldsymbol{E} \cdot \boldsymbol{r}}{\mathrm{i}\hbar\omega} = \frac{q}{\hbar} \boldsymbol{A}' \cdot \boldsymbol{r} \tag{2.6}$$

from equation (2.2) we obtain the new vector and new scalar potentials for the $\boldsymbol{p} \cdot \boldsymbol{A}$ gauge as

$$\boldsymbol{A}' = \frac{\boldsymbol{E}}{\mathrm{i}\omega} \qquad \phi' = 0. \tag{2.7}$$

The connection between the old and new wavefunction is determined by equation (2.3).

In perturbation schemes for studying the optical response, conventionally people use exactly the same set of unperturbed wavefunctions $\psi_n^0(\boldsymbol{r}, t)$ of the Hamiltonian $\hat{H}_0$ (when $\boldsymbol{A} = 0$ and $\phi = 0$ in equation (2.1)) to serve as the expansion basis for both $\boldsymbol{E} \cdot \boldsymbol{r}$ and $\boldsymbol{p} \cdot \boldsymbol{A}$ gauges [3, 5, 8]. However, we should point out that the wavefunctions for both $\boldsymbol{E} \cdot \boldsymbol{r}$ and $\boldsymbol{p} \cdot \boldsymbol{A}$ gauges (before and after gauge transformation) should also be restricted by the gauge phase factor $F_g$ from equation (2.3); therefore the two basis sets for the two gauges are *not* exactly the same unperturbed wavefunctions $\psi_n^0(\boldsymbol{r}, t)$—they differ by the gauge phase factor $F_g$. And the Hamiltonians under the two gauges ($\boldsymbol{E} \cdot \boldsymbol{r}$ and $\boldsymbol{p} \cdot \boldsymbol{A}$) are not necessarily equivalent if they are treated independently and are isolated from the connection between the wavefunctions under the two gauges. Unfortunately, this crucial point has not been clearly stated, and was obviously missed in the perturbation scheme studies [3, 5, 6]. In the current–current correlation scheme, in particular, the gauge phase factor's contribution is obviously ignored and the $A^2(t)$-term is considered as having no physical meaning†. Thus the current–current correlation is conventionally reduced to a $J_0 J_0$-formula such as equation (1.2), and the equivalence between current–current and dipole–dipole correlations is usually considered as represented by $J_0 J_0$- and $DD$-correlations with exactly the same basis of unperturbed wavefunctions [1, 3, 5, 8, 11, 15].

An elegant review by Langhoff, Epstein, and Karplus covered the topics of time-dependent perturbation theory [36]; they firmly pointed out that the time-dependent phase in a wavefunction is very much essential and that the incorrect treatment of the time-dependent phase will cause secular divergence in time-dependent perturbations. In field theory, it is also well understood that inappropriate treatment of the phase factor will cause divergence [2]. Since the gauge phase factor, equation (2.6), is obviously time dependent, neglecting this phase factor will cause a ZFD in the susceptibility computations, as in the examples that we will provide in the following sections.

† On p 111 of Butcher and Cotter's book [3], they directly pointed out that the $\boldsymbol{A}^2(t)$ term has no physical effects and merely introduces a time-dependent phase factor into the wavefunction. Thus, for $n \geqslant 2$, obviously equation (1.4) and equation (1.2) simply effect the following $\hat{\boldsymbol{D}}$–$\hat{\boldsymbol{J}}$ replacement: $\hat{\boldsymbol{D}} \rightarrow \hat{\boldsymbol{J}}/(\mathrm{i}\omega_i)$. (See p 107 of [3].) Although in Bloembergen's notes [5] $\boldsymbol{A}^2(t)$ is seemingly included in the $\chi^{(2)}$-derivations (p 35), the results obtained by Bloembergen (equation (2-48)) are exactly the same as the results obtained without considering the $\boldsymbol{A}^2(t)$ terms. (See equation (4.61) given by Butcher and Cotter [3] with the $\hat{\boldsymbol{D}}$–$\hat{\boldsymbol{J}}$ replacement.) This shows that $\chi^{(2)}$-computation is a special case in non-linear susceptibility [11]. And in the $\chi^{(3)}$-derivations on p 172 of [5], the expansion basis is the unperturbed wavefunctions.

## 3. Linear response from dipole–dipole correlation

Having appreciated the importance of the gauge phase factor, we choose the following single-electron periodic models—the SSH (or Hückel) model and the TLM model—as our examples for the following reasons:

(i) These periodic models were widely applied in polymer theory in the 1980s and early 1990s; remarkable results have been obtained [37].
(ii) The optical susceptibilities obtained from these models are analytically solvable and can be compared with previous results [17, 20].
(iii) In both the SSH and the TLM models, Peierls instability [37] leads to the semiconductivity of the two-band structure—with the valence band fully filled and the conduction band empty. It is very obvious from the physical point of view that, as the frequency of the electrical field goes 0 (reaches the static electric field), the linear conductivity $\sigma^{(1)}$ and the linear susceptibility $\chi^{(1)}$ will not be reduced to the Drude formula as in the case of metals [12, 13], and will not cause a ZFD problem.

In this section, we will first discuss $\chi^{(1)}$ and $\sigma^{(1)}$ for infinite chains (where the number of (CH) units $N$ goes to infinity) under the $DD$-correlation.

### *3.1. The SSH Hamiltonian*

Based on the periodic TBA, the SSH Hamiltonian [21] is given by

$$H_{SSH} = -\sum_{l,s}\left[t_0 + (-1)^l\frac{\Delta}{2}\right](\hat{C}^{\dagger}_{l+1,s}\hat{C}_{l,s} + \hat{C}^{\dagger}_{l,s}\hat{C}_{l+1,s}) \tag{3.1}$$

where $t_0$ is the transfer integral for the nearest-neighbour sites, $\Delta$ is the gap parameter, and $\hat{C}^{\dagger}_{l,s}$ ($\hat{C}_{l,s}$) creates (annihilates) a $\pi$-electron at site $l$ with spin $s$. In the continuum limit, the above SSH model will give the TLM model [22]. For the SSH model, each site is occupied by one electron.

If we want to include the electron–photon interaction $\boldsymbol{E}\cdot\boldsymbol{r}$ directly via the polarization operator $\hat{\boldsymbol{P}}$, where

$$\hat{P} = \sum_l R_l\hat{C}^{\dagger}_l\hat{C}_l \tag{3.2}$$

and

$$R_l = la + (-1)^l u \tag{3.3}$$

is the site-$l$ position with the lattice constant $a$ and dimerized constant $u$ [21], we will face the problem of ill definition of $\hat{\boldsymbol{P}}$ in periodic systems [11, 15, 20]. To solve this problem, we consider an imposed periodic condition for the position operator $\boldsymbol{r}$ [14, 15]. Expressing the position operator $\boldsymbol{r}$ in terms of Bloch states $|n,\boldsymbol{k}\rangle = u_{n,\boldsymbol{k}}(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$, where $u_{n,\boldsymbol{k}}(\boldsymbol{r})$ is a function that is periodic under translation by a lattice vector [14], we will be able to satisfy the periodic condition for $\boldsymbol{r}$ as follows:

$$\boldsymbol{r}_{n\boldsymbol{k},n'\boldsymbol{k}'} = \mathrm{i}\delta_{n,n'}\nabla_k\delta(\boldsymbol{k}-\boldsymbol{k}') + \Omega_{n,n'}(\boldsymbol{k})\delta(\boldsymbol{k}-\boldsymbol{k}') \tag{3.4}$$

and

$$\Omega_{n,n'}(\boldsymbol{k}) = \frac{\mathrm{i}}{v}\int_v u^{*}_{n,\boldsymbol{k}}(\boldsymbol{r})\,\nabla_k u_{n',\boldsymbol{k}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \tag{3.5}$$

where $v$ is the unit-cell volume.

We diagonalize the Hamiltonian equation (3.1) in the momentum space by applying consecutive transformations under the operators $\hat{C}_l^\dagger$ and $\hat{C}_l$ [21]. The excitations of electrons in the conduction band and valence band with momentum $k$ and spin $s$, $\hat{a}_{k,s}^{\dagger c}(t)$ and $\hat{a}_{k,s}^{\dagger v}(t)$, respectively, are obtained.

If we choose the spinor description $\hat{\psi}_{k,s}^\dagger(t) = (\hat{a}_{k,s}^{\dagger c}(t),\ \hat{a}_{k,s}^{\dagger v}(t))$, the SSH Hamiltonian including $\boldsymbol{E}\cdot\boldsymbol{r}$ in momentum space is described by

$$\hat{H}_{SSH}(k,t) = \hat{H}_0 + \hat{H}_{\boldsymbol{E}\cdot\boldsymbol{r}} \tag{3.6}$$

where

$$\hat{H}_0 = \sum_{-\pi/(2a)\leqslant k\leqslant \pi/(2a),s} \varepsilon(k)\hat{\psi}_{k,s}^\dagger(t)\sigma_3\hat{\psi}_{k,s}(t) \tag{3.7}$$

and

$$\hat{H}_{\boldsymbol{E}\cdot\boldsymbol{r}} = -\hat{D}E_0 \mathrm{e}^{\mathrm{i}\omega t}. \tag{3.8}$$

From equation (3.4), the dipole operator $\hat{D}$ can be obtained as follows:

$$\hat{D} = e\sum_{-\pi/(2a)\leqslant k\leqslant \pi/(2a),s}\left(\beta(k)\ \hat{\psi}_{k,s}^\dagger\sigma_2\hat{\psi}_{k,s} + \mathrm{i}\,\frac{\partial}{\partial k}\hat{\psi}_{k,s}^\dagger\hat{\psi}_{k,s}\right) \tag{3.9}$$

where

$$\beta(k) = -\frac{\Delta t_0 a}{\varepsilon^2(k)} + u \tag{3.10}$$

is the coefficient related to the interband transition between the conduction and valence bands in a unit cell of size $2a$ and the second term in equation (3.9) is related to the intraband transition [11], $e$ is the electric charge, and the $\vec{\sigma}$ are the Pauli matrices. $u$ is a dimerized constant related to the lattice distortion [21].

### *3.2. Linear response through the $\boldsymbol{E}\cdot\boldsymbol{r}$ gauge*

For the linear susceptibility, $\chi_{SSH}^{(1)}(\Omega,\omega_1)$ can be obtained from equation (1.4) and equation (3.9):

$$\begin{aligned}\chi_{SSH}^{(1)}(-\omega_1,\omega_1) = 2\left[\frac{\mathrm{i}}{\hbar}\right]e^2\frac{1}{L}\sum_k\int_{-\infty}^{\infty}\mathrm{Tr}\Bigg\{&\mathrm{i}\,\frac{\partial}{\partial k}\left[G(k,\omega)\,\mathrm{i}\,\frac{\partial}{\partial k}\left[G(k,\omega-\omega_1)\right]\right]\\ &+\beta(k)\sigma_2 G(k,\omega)\,\mathrm{i}\,\frac{\partial}{\partial k}\left[G(k,\omega-\omega_1)\right] + \mathrm{i}\,\frac{\partial}{\partial k}\left[\beta(k)G(k,\omega)\sigma_2 G(k,\omega-\omega_1)\right]\\ &+\beta(k)\sigma_2 G(k,\omega)\beta(k)\sigma_2 G(k,\omega-\omega_1)\Bigg\}\frac{\mathrm{d}\omega}{2\pi}\end{aligned} \tag{3.11}$$

where $L$ is the total chain length and the Green function $G(k,\omega)$ is the Fourier transform of $G(t-t') \equiv -\mathrm{i}\langle\hat{T}\hat{\psi}(t)\hat{\psi}^\dagger(t')\rangle$, given as follows:

$$G(k,\omega) = \frac{\omega+\omega_k\sigma_3}{\omega^2-\omega_k^2+\mathrm{i}\epsilon} \tag{3.12}$$

with $\omega_k \equiv \varepsilon(k)/\hbar$ and $\epsilon \equiv 0^+$.

From equation (3.11), we have $\chi_{SSH}^{(1)}(\omega) \equiv \chi_{SSH}^{(1)}(-\omega,\omega)$:

$$\chi_{SSH}^{(1)}(\omega) = \frac{e^2(2t_0a)}{2\pi\Delta^2}\int_1^{1/\delta}\frac{(1-\eta\delta x^2)^2\,\mathrm{d}x}{[(1-\delta^2x^2)(x^2-1)]^{1/2}x^2(x^2-z^2)} \tag{3.13}$$

where $x \equiv \hbar\omega_k/\Delta$, $z \equiv \hbar\omega/(2\Delta)$, $\delta \equiv \Delta/(2t_0)$, and the relative distortion $\eta \equiv (2u)/a$. Equation (3.13) can be integrated numerically if one changes $x$ to $x + \mathrm{i}\epsilon$ when considering the lifetime of the state [26–29]. For polyacetylene, by choosing $t_0 = 2.5$ eV, $\Delta = 0.9$ eV, $a = 1.22$ Å, $u = 0.04$ Å, and $\epsilon \sim 0.03$ [26–29], we obtain $\delta = 0.18$ and $\eta = 0.07$. The values of $|\chi^{(1)}_{SSH}|$ with and without the $\eta$-contribution are plotted in figure 1. As we can clearly see from the graph, the contribution of the relative distortion $\eta$ is very small (about 1%). We can observe the 'Umklapp-enhancement' peak at $z = 1$ compared with the peak $z = 5.56$ $(1/\delta)$. These results were also discussed in previous work [20].

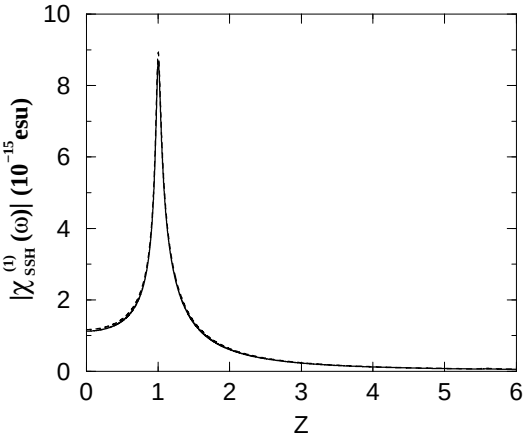


**Figure 1.** $|\chi^{(1)}_{SSH}(\omega)|$ with $z \equiv \hbar\omega/(2\Delta)$, $\epsilon = 0.03$, $\delta = 0.18$, and for $\eta = 0.07$ (solid line) or $\eta = 0$ (dashed line).

If the continuum limit is applied—that is, $\delta \to 0^+$, $\eta \to 0^+$, $\epsilon \to 0^+$, and $2t_0 a \to \hbar v_F$—then the above integral equation (3.13) approaches the linear optical susceptibility $\chi^{(1)}_{TLM}(\omega)$ in the TLM model [22] as follows:

$$\chi^{(1)}_{TLM}(\omega) = -\frac{e^2\hbar v_F}{2\pi\Delta^2 z^2}(1 - f(z)) \tag{3.14}$$

where

$$f(z) \equiv \begin{cases} \dfrac{\arcsin(z)}{z\sqrt{1-z^2}} & z^2 < 1 \\ -\dfrac{\cosh^{-1}(z)}{z\sqrt{z^2-1}} + \dfrac{\mathrm{i}\pi}{2z\sqrt{z^2-1}} & z^2 > 1. \end{cases} \tag{3.15}$$

The conductivity $\sigma(\omega)$ is given by $-\mathrm{i}\omega\chi^{(1)}$, and $\mathrm{Re}[\sigma^{(1)}(\omega)]$ is exactly the same as previous results [17, 20].

The calculated $\chi^{(1)}_{TLM}$ and the absolute value of $\chi^{(1)}_{TLM}$ are shown in figure 2 and figure 3, respectively.

Indeed, the above computations are based on the perturbation scheme with $\hat{\psi}^{\dagger}_{k,s}(t)$ and $\hat{\psi}_{k,s}(t)$ as the unperturbed creation and annihilation operators under the Hamiltonian $\hat{H}_0$ (equation (3.7)). From the above figures and expressions, it is very obvious that the $DD$-correlation (or $\boldsymbol{E}\cdot\boldsymbol{r}$) approach will not have a ZFD in $\chi^{(1)}$ or $\sigma^{(1)}$. Straightforward computations readily show that equation (3.13) and equation (3.14) obey the KK relations. These results are certainly reasonable in view of the physical picture.

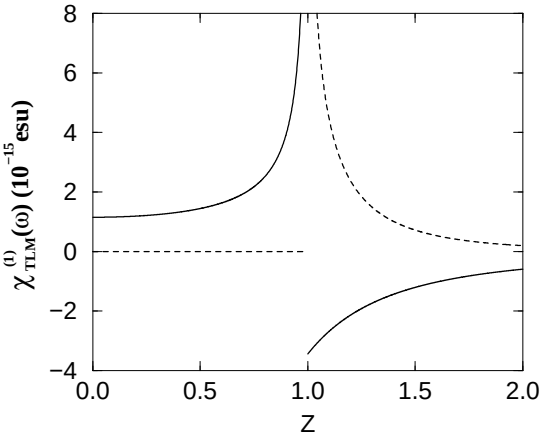


**Figure 2.** The real part (solid line) and the imaginary part (dashed line) of $\chi^{(1)}_{TLM}(\omega)$ with $z \equiv \hbar\omega/(2\Delta)$.

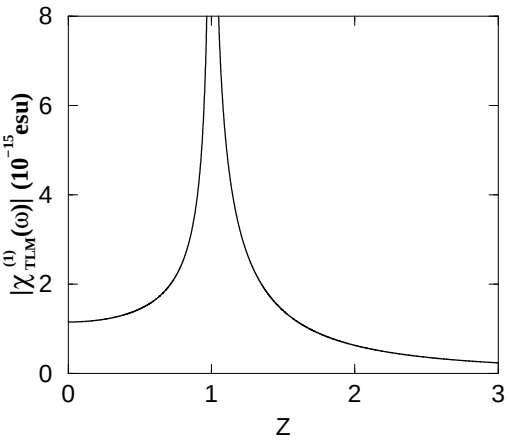


**Figure 3.** $|\chi^{(1)}_{TLM}(\omega)|$ with $z \equiv \hbar\omega/(2\Delta)$.

## 4. Linear response through current–current correlation

### *4.1. The SSH Hamiltonian in the vector potential form*

The tight-binding Hamiltonian in the $\boldsymbol{p}\cdot\boldsymbol{A}$ form should be invariant under gauge transformation (equation (2.2) and equation (2.3)) [20, 23]. If we change the phase of the one-particle wavefunction in the tight-binding approximation:

$$\hat{C}'_s(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\theta}\hat{C}_s(\boldsymbol{r}) \tag{4.1}$$

we must modify the kinetic energy term. The unperturbed Hamiltonian $\hat{H}_0$ in the Wannier-function basis as follows:

$$H_0 \equiv \sum_{s,\boldsymbol{r},\boldsymbol{r}'} t(\boldsymbol{r}-\boldsymbol{r}')(C^{\dagger}_s(\boldsymbol{r})C_s(\boldsymbol{r}') + C^{\dagger}_s(\boldsymbol{r}')C_s(\boldsymbol{r})) \tag{4.2}$$

is modified as

$$H_0 \to H'_0 = \sum_{s,\boldsymbol{r},\boldsymbol{r}'} t(\boldsymbol{r}-\boldsymbol{r}')\left[C'^{\dagger}_s(\boldsymbol{r})\exp\left(-\mathrm{i}q\int_{\boldsymbol{r}}^{\boldsymbol{r}'}\mathrm{d}\boldsymbol{x}\cdot\boldsymbol{A}(x)/\hbar\right)C'_s(\boldsymbol{r}')\right.$$
$$\left. + C'^{\dagger}_s(\boldsymbol{r}')\exp\left(\mathrm{i}q\int_{\boldsymbol{r}}^{\boldsymbol{r}'}\mathrm{d}\boldsymbol{x}\cdot\boldsymbol{A}(x)/\hbar\right)C'_s(\boldsymbol{r})\right] \tag{4.3}$$

where $t(\boldsymbol{r}-\boldsymbol{r}')$ represents the hopping from position $\boldsymbol{r}$ to $\boldsymbol{r}'$, $C_s^\dagger(\boldsymbol{r})$ creates an electron at site $\boldsymbol{r}$ with spin $s$, $q$ is the particle charge. The above transformation is also known as the Peierls substitution [20]. Equation (4.3) has some kind of general meaning in TBA models and is frequently applied in theoretical work [38].

If the function $f(\boldsymbol{r},t)$ is arbitrary in equation (2.2), it is easy to verify that the above TBA Hamiltonian equation (4.3) is invariant if the local phase $\theta$ in equation (4.1) is defined as

$$\theta(\boldsymbol{r},t) \equiv \frac{q}{\hbar} f(\boldsymbol{r},t) \equiv \frac{q}{\hbar}\boldsymbol{A}\cdot\boldsymbol{r}. \tag{4.4}$$

As a specific example, the SSH Hamiltonian with the vector potential $A$ should be as follows (we change $\boldsymbol{A}$ to $A$ since it is the 1D case and $e$ is the electron charge)†:

$$H_{SSH}(A) = -\sum_{l,s}\left[t_0 + (-1)^l\frac{\Delta}{2}\right](\hat{C}'^{\dagger}_{l+1,s}\mathrm{e}^{-\mathrm{i}eA(R_l-R_{l+1})/\hbar}\hat{C}'_{l,s} + \hat{C}'^{\dagger}_{l,s}\mathrm{e}^{\mathrm{i}eA(R_l-R_{l+1})/\hbar}\hat{C}'_{l+1,s}). \tag{4.5}$$

Equation (4.5) has been derived in full detail by Gebhard *et al* [20].

### *4.2. The result without the gauge phase factor*

Under the assumption that the gauge phase factor (4.4) has no physical meaning [3], we ignore the phase $\theta$ and treat the creation and annihilation operators $\{\hat{C}'^{\dagger}_l\}$ and $\{\hat{C}'_l\}$ as the same as the unperturbed creation and annihilation operators $\{\hat{C}^\dagger_l\}$ and $\{\hat{C}_l\}$ [20, 23] defined in equation (3.1).

The above $\boldsymbol{p}\cdot\boldsymbol{A}$ Hamiltonian equation (4.5) can be expanded in powers of the external vector potential $A$, and we obtain the following:

$$\hat{H}_{SSH}(A) = \hat{H}_0 - \hat{J}_0 A + \mathrm{O}(A^2) \tag{4.6}$$

where $H_0$ is given by equation (3.1), and

$$\hat{J}_0 = -\sum_{l,s}\mathrm{i}\frac{e}{\hbar}\left[t_0 + (-1)^l\frac{\Delta}{2}\right]\left[a - 2(-1)^l u\right](\hat{C}^\dagger_{l+1,s}\hat{C}_{l,s} - \hat{C}^\dagger_{l,s}\hat{C}_{l+1,s}). \tag{4.7}$$

The current operator $\hat{J}$ is obtained from the following equation:

$$\hat{J} = \frac{\mathrm{i}}{\hbar}[\hat{P},\hat{H}]. \tag{4.8}$$

From equation (3.2) and equation (4.6), we obtain the current operator for the SSH Hamiltonian as follows:

$$\hat{J}_{SSH} = \hat{J}_0 + \hat{J}_1 A \tag{4.9}$$

where $\hat{J}_0$ is defined by equation (4.7), and $\hat{J}_1$ is defined as follows:

$$\hat{J}_1 = -\sum_{l,s}\left(\frac{e}{\hbar}\right)^2\left[t_0 + (-1)^l\frac{\Delta}{2}\right]\left[a - 2(-1)^l u\right]^2(\hat{C}^\dagger_{l+1,s}\hat{C}_{l,s} + \hat{C}^\dagger_{l,s}\hat{C}_{l+1,s}). \tag{4.10}$$

As in the computations in section 3, we transform the Hamiltonian equation (4.6) and the current operators equation (4.7) and equation (4.10) into the momentum space and obtain the following:

$$\hat{J}_0(k) = \frac{ea}{\hbar}\sum_{k,s}\left[A_0(k)\hat{\psi}^\dagger_{k,s}(t)\sigma_3\hat{\psi}_{k,s}(t) + B_0(k)\hat{\psi}^\dagger_{k,s}(t)\sigma_1\hat{\psi}_{k,s}(t)\right] \tag{4.11}$$

† Please bear in mind that we have not written out another term—the scalar potential $\phi(R_l,t)$, which is time dependent in the above Hamiltonian equation (4.5). However, in section 4 (the $\boldsymbol{p}\cdot\boldsymbol{A}$ case), that term is 0. The $\phi(R_l,t)$ term will appear if the gauge transformation is made.

and

$$\hat{J}_1(k) = \left(\frac{ea}{\hbar}\right)^2 \sum_{k,s} \left[ A_1(k)\hat{\psi}^{\dagger}_{k,s}(t)\sigma_3\hat{\psi}_{k,s}(t) + B_1(k)\hat{\psi}^{\dagger}_{k,s}(t)\sigma_1\hat{\psi}_{k,s}(t) \right] \quad (4.12)$$

with $A_0(k)$, $B_0(k)$, $A_1(k)$, and $B_1(k)$ defined as follows:

$$\begin{aligned} A_0(k) &= -\frac{(2t_0)^2(1-\delta^2)\sin(2ka)}{2\varepsilon(k)} \\ B_0(k) &= -\frac{2t_0\Delta}{\varepsilon(k)} + \eta\varepsilon(k) \end{aligned} \quad (4.13)$$

and

$$\begin{aligned} A_1(k) &= (1+\eta^2)\varepsilon(k) - \frac{4t_0\Delta\eta}{\varepsilon(k)} \\ B_1(k) &= \frac{(2t_0)^2\eta(1-\delta^2)\sin(2ka)}{\varepsilon(k)}. \end{aligned} \quad (4.14)$$

$\eta$, $\delta$, and $\varepsilon(k)$ in the above equations are defined the same way as in section 3.

Applying equation (1.2) and equation (1.3), and considering the diamagnetic current $\hat{J}_1(k)A$ obtained from the above SSH Hamiltonian equation (4.6), we obtain the linear susceptibility under $J_0J_0$-correlation as follows:

$$\chi'^{(1)}_{SSH}(-\omega_1,\omega_1) = \frac{\chi'^{(1)}_{j_0j_0}(-\omega_1,\omega_1)}{-\mathrm{i}\omega_1^2} + \frac{\chi'^{(1)}_{j_1}(-\omega_1,\omega_1)}{-\mathrm{i}\omega_1^2} \quad (4.15)$$

with

$$\begin{aligned} \chi'^{(1)}_{j_0j_0}(-\omega_1,\omega_1) = 2\left[\frac{1}{\hbar}\right]\left(\frac{ea}{\hbar}\right)^2 \frac{1}{L}\sum_k \int_{-\infty}^{\infty} \mathrm{Tr}\{&A_0(k)\sigma_3 G(k,\omega)A_0(k)\sigma_3 G(k,\omega-\omega_1) \\ &+ A_0(k)\sigma_3 G(k,\omega)B_0(k)\sigma_1 G(k,\omega-\omega_1) \\ &+ B_0(k)\sigma_1 G(k,\omega)A_0(k)\sigma_3 G(k,\omega-\omega_1) \\ &+ B_0(k)\sigma_1 G(k,\omega)B_0(k)\sigma_1 G(k,\omega-\omega_1)\}\frac{\mathrm{d}\omega}{2\pi} \end{aligned} \quad (4.16)$$

$$\begin{aligned} \chi'^{(1)}_{j_1}(-\omega_1,\omega_1) = -2\left(\frac{ea}{\hbar}\right)^2 \frac{1}{L}\sum_k \int_{-\infty}^{\infty} \mathrm{Tr}\{&A_1(k)\sigma_3 G(k,\omega-\omega_1) \\ &+ B_1(k)\sigma_1 G(k,\omega-\omega_1)\}\frac{\mathrm{d}\omega}{2\pi} \end{aligned} \quad (4.17)$$

where the Green function $G(k,\omega)$ is defined by equation (3.12). By straightforward computation, we can obtain the following:

$$\chi'^{(1)}_{j_0j_0}(\omega) = -2\mathrm{i}\frac{e^2(2t_0a)}{\pi\hbar^2}\int_1^{1/\delta} \frac{(1-\eta\delta x^2)^2\,\mathrm{d}x}{[(1-\delta^2x^2)(x^2-1)]^{1/2}(x^2-z^2)} \quad (4.18)$$

and

$$\chi'^{(1)}_{j_1}(\omega) = 2\mathrm{i}\frac{e^2(2t_0a)}{\pi\hbar^2}\int_1^{1/\delta} \frac{[(1+\eta^2)\delta^2x^2 - 2\eta\delta]\,\mathrm{d}x}{[(1-\delta^2x^2)(x^2-1)]^{1/2}}. \quad (4.19)$$

Equation (4.19) is a constant independent of $z$ and it is obviously not the term $\langle[j_0,\, j_0]\rangle(\omega=0)$ suggested by Batistic and Bishop [19].

From equation (4.15), equation (4.18), and equation (4.19), we obtain

$$\chi'^{(1)}_{SSH}(\omega) = \frac{e^2(2t_0a)}{2\pi\Delta^2 z^2}\left\{\int_1^{1/\delta}\frac{(1-\eta\delta x^2)^2\,\mathrm{d}x}{[(1-\delta^2x^2)(x^2-1)]^{1/2}(x^2-z^2)} - \int_1^{1/\delta}\frac{[(1+\eta^2)\delta^2x^2-2\eta\delta]\,\mathrm{d}x}{[(1-\delta^2x^2)(x^2-1)]^{1/2}}\right\} \tag{4.20}$$

where $x$, $\Delta$, $z$, $\delta$, and $\eta$ are all the same as in equation (3.13).

If the continuum limit is applied (as in section 3), we find that the contribution (equation (4.19)) from the diamagnetic term (the $J_1$-term) disappears. We obtain the following susceptibility under TLM models:

$$\chi'^{(1)}_{TLM}(\omega) = \frac{e^2\hbar v_F}{2\pi\Delta^2 z^2}f(z) \tag{4.21}$$

where $f(z)$ is defined by equation (3.15).

We plot $|\chi'^{(1)}_{SSH}|$, $\chi'^{(1)}_{TLM}$, and $|\chi'^{(1)}_{TLM}|$ in figure 4, figure 5, and figure 6, respectively, with the same parameters as in section 3. From the analytical form equation (4.21) and the figures, we find that the first term in equation (3.14) disappears in this $J_0J_0$-formula; however, this important feature has not been reported in the previous works [17, 20].

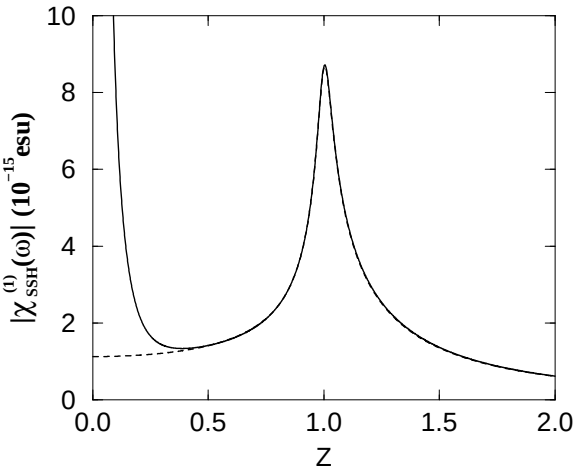


**Figure 4.** $|\chi'^{(1)}_{SSH}(\omega)|$ obtained through $J_0J_0$-correlation (solid line) versus $|\chi^{(1)}_{SSH}(\omega)|$ obtained through $DD$-correlation (dashed line) with $z \equiv \hbar\omega/(2\Delta)$, for $\epsilon = 0.03$, $\eta = 0.07$, and $\delta = 0.18$. $|\chi'^{(1)}_{SSH}(\omega)|$ obviously shows the ZFD if the gauge phase factor is not considered.

Within TLM models, the missing diamagnetic term can be understood directly from the Lagrangian†, which only contains the first-order terms of the momentum operator $\hat{p}$. Obviously, equation (4.20) and equation (4.21) strongly diverge when $z \to 0$ for the real parts in both equation (4.20) and equation (4.21). The above results are certainly wrong, since they do not follow the KK relations—besides the ZFD problem. Careful comparisons were made between equations (4.20), (4.21) and equations (3.13), (3.14); it was found that the correct imaginary parts obtained from the $J_0J_0$-correlation equation (1.2) and equation (1.3) are still maintained. The comparisons also show that the absorption part (related to imaginary part) based on the $J_0J_0$-correlation will still be correct. This conclusion is not novel to those working on the transport problem: that the Kubo formula based on $J_0J_0$-correlation can be applied is

† The Lagrangian for the TLM models is described as [1, 22] $L = \int \mathrm{d}x\ \hat{\psi}^\dagger(x)\{\mathrm{i}\hbar\,\partial_t + \mathrm{i}\sigma_3 v_F\,\partial_x + \sigma_1\Delta\}\hat{\psi}(x)$. The vector potential $A$ is included by changing $-\mathrm{i}\hbar\,\partial_x$ to $\mathrm{i}\hbar\,\partial_x - eA$; it does not contain the $A^2$-term from the Lagrangian directly.

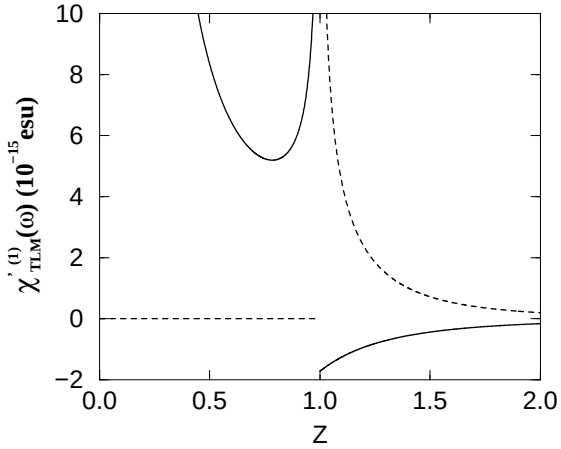


**Figure 5.** The real part (solid line) and the imaginary part (dashed line) of $\chi'^{(1)}_{TLM}(\omega)$ with $z \equiv \hbar\omega/(2\Delta)$. This shows the ZFD that is present in the real part when the gauge phase factor is not considered.

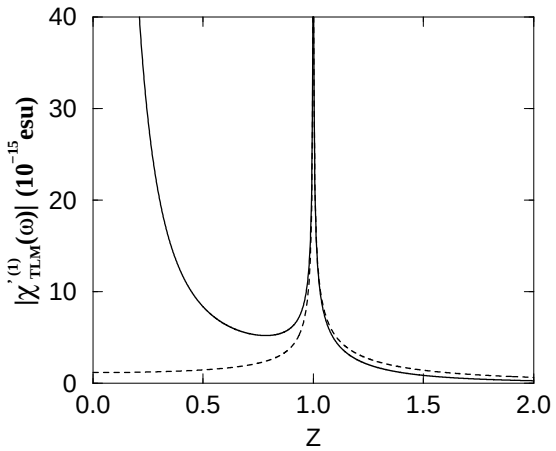


**Figure 6.** $|\chi'^{(1)}_{TLM}(\omega)|$ obtained through the $J_0J_0$-correlation (solid line) compared with $|\chi^{(1)}_{TLM}(\omega)|$ obtained through the $DD$-correlation or $JJ$-correlation, considering the gauge phase factor (dashed line) with $z \equiv \hbar\omega/(2\Delta)$. $|\chi'^{(1)}_{TLM}(\omega)|$ shows a ZFD.

already common knowledge [2]. People usually resort to the KK relation to avoid the real-part ZFD difficulty [12].

But how to explain these difficulties (such as the ZFD and violation of the KK relation in equation (4.20) and equation (4.21)) now becomes a task. The long-wavelength approximation used in the above examples already eliminates the possible ZFD caused by the limit sequence for $\boldsymbol{k}$ and $\omega$†. One might argue that the ZFD problem could be solved by including a diamagnetic term, such as the effective mass $m^*$, and that the electron-density ($n_0$) assumption can be made in this solid-state problem [13, 14]. In fact, these assumptions are awkward, because the parameters $m^*$ and $n_0$ in diamagnetic terms cannot be predicted in the models. For example,

† Mahan and Haug suggest a possible sequence limit [2, 10]: firstly setting the wavevector $\boldsymbol{k} \to 0$, and secondly setting the frequency $\omega \to 0$ when considering the linear conductivity $\sigma^{(1)}$ at zero frequency. If the long-wavelength limit applies, the sequence condition for $\boldsymbol{k}$ and $\omega$ is automatically satisfied.

in the above SSH and TLM models, $m^*$ and $n_0$ could be arbitrary [37]. In the TLM model, you cannot include the diamagnetic term directly from the Lagrangian†. For the SSH model, the diamagnetic term can be included in equation (4.9) from the TBA Hamiltonian equation (4.6), but you still cannot solve this ZFD problem, as we clearly see from figure 4 and equation (4.20). Moreover, if one uses diamagnetic terms to cancel the ZFD, the assumption regarding this property of the medium should be included in the proof [2, 9, 10].

The conventional way to treat this ZFD problem is to separate the divergent terms from the convergent term and discard them, with some possible physical explanations [11]. Fortunately, in linear response theory, neglect of the diamagnetic term does not cause so much trouble, since the imaginary part of the $J_0J_0$-correlation is still correct [2, 10, 12]. It is not a big surprise that people have not taken this ZFD problem seriously, although doubts have always existed regarding the two gauges, especially in models [8, 15, 35].

### *4.3. Resolving the ZFD by using the gauge phase factor*

As we demonstrated above, the diamagnetic term directly obtained from equation (4.5) cannot solve the ZFD problem. This ZFD problem, which is a conceptual problem as we have already pointed out in section 2, is caused by the conventional careless treatment of the gauge phase factor in optical response theory. In this part, after considering the contribution of the gauge phase factor in the $JJ$-correlation, we will recover the same results for $\chi^{(1)}$ as were obtained from the $DD$-correlation within both the SSH and TLM models.

As we discussed in section 2, the new creation operator $\hat{C}_l'^{\dagger}$ and annihilation operator $\hat{C}_l'$ in equation (4.5) should be differentiated by the gauge phase factor from the unperturbed creation operator $\hat{C}_l^{\dagger}$ and annihilation operators $\hat{C}_l$. Following equation (2.3), we obtain the following, after the local phase factor is considered‡:

$$\hat{C}_l' = \mathrm{e}^{\mathrm{i}eAR_l/\hbar}\hat{C}_l. \tag{4.22}$$

Thus, by the above relation, the Hamiltonian in momentum space should undergo the following change:

$$\hat{H}(k) \to \hat{H}(\kappa) \tag{4.23}$$

where

$$\kappa = k + \frac{eA}{\hbar}. \tag{4.24}$$

For the SSH Hamiltonian, we have the following new Hamiltonian:

$$H^{new}(k) = H_0^{new}(k) + H_1^{new}(k)A + \mathrm{O}(A^2) \tag{4.25}$$

where $H_0^{new}(k)$ is defined as in equation (3.7) and

$$H_1^{new}(k) = -\frac{ea}{\hbar}\sum_k B_0(k)\hat{\psi}_{k,s}^{\dagger}(t)\sigma_1\hat{\psi}_{k,s}(t). \tag{4.26}$$

The new current operator $\hat{J}^{new}(k)$ can be obtained from the commutator equation $[\hat{D}(k), \hat{H}^{new}(k)]/(\mathrm{i}\hbar)$ as follows:

$$\hat{J}^{new}(k) = \hat{J}_0^{new}(k) + \hat{J}_1^{new}(k)A + \mathrm{O}(A^2) \tag{4.27}$$

† The Lagrangian for the TLM models is described as [1, 22] $L = \int \mathrm{d}x\ \hat{\psi}^{\dagger}(x)\{\mathrm{i}\hbar\,\partial_t + \mathrm{i}\sigma_3 v_F\,\partial_x + \sigma_1\Delta\}\hat{\psi}(x)$. The vector potential $A$ is included by changing $-\mathrm{i}\hbar\partial_x$ to $\mathrm{i}\hbar\,\partial_x - eA$; it does not contain the $A^2$-term from the Lagrangian directly.

‡ Strictly speaking, this gauge phase factor in this paper is an approximate of the diagonal elements in the matrix of the dipole operator $\hat{D}$ [20, 23]. As in the example we give above, the diagonal term is enough for obtaining convergent results in the linear case. But for non-linear cases like $\chi^{(2)}$ and $\chi^{(3)}$, the non-diagonal terms of the $\hat{D}$-matrix will be important for cancelling the ZFD problem through the $JJ$-correlation. Certainly the procedures for including the full gauge phase factor will be complicated and thus impractical for the $JJ$-correlation.

where $\hat{J}_0^{new}(k)$ is exactly the same as equation (4.11) and

$$\hat{J}_1^{new}(k) = \left(\frac{ea}{\hbar}\right)^2 \sum_{k,s} \left[ A_1^{new}(k)\hat{\psi}_{k,s}^{\dagger}(t)\sigma_3\hat{\psi}_{k,s}(t) + B_1^{new}(k)\hat{\psi}_{k,s}^{\dagger}(t)\sigma_1\hat{\psi}_{k,s}(t) \right] \tag{4.28}$$

where

$$\begin{aligned} A_1^{new}(k) &= \frac{B_0^2(k)}{\varepsilon(k)} \\ B_1^{new}(k) &= -A_0(k)\left[\frac{2t_0\Delta}{\varepsilon^2(k)} + \eta\right]. \end{aligned} \tag{4.29}$$

$A_0(k)$ and $B_0(k)$ are defined in equation (4.13).

After considering the gauge factor, we find that the new current operator $J_0^{new}(k)$ is the same as the static current $J_0(k)$, but that $J_1^{new}(k)$ is different from the static current $J_1(k)$; we call the current differences between $J^{new}(k)$ and the static currents $J_0(k)$, $J_1(k)$, etc induced field currents (IFC), since they are introduced by the gauge field.

Through the evolution operator in the interaction picture [2], it is easy to derive the formula for the $\chi_{jj}^{(1)}$ as follows:

$$\chi_{j^{new}j^{new}}^{(1)}(-\omega_1, \omega_1) = \frac{\chi_{j_0^{new}h_1^{new}}^{(1)}(-\omega_1, \omega_1) + \chi_{j_1^{new}}^{(1)}(-\omega_1, \omega_1)}{-\mathrm{i}\omega_1^2} \tag{4.30}$$

where

$$\begin{aligned} \chi_{j_0^{new}h_1^{new}}^{(1)}(-\omega_1, \omega_1) = 2\left[\frac{1}{\hbar}\right]\left(\frac{ea}{\hbar}\right)^2 \frac{1}{L}\sum_k \int_{-\infty}^{\infty} \mathrm{Tr}\{&A_0(k)\sigma_3 G(k,\omega)B_0(k)\sigma_1 G(k,\omega-\omega_1) \\ &+ B_0(k)\sigma_1 G(k,\omega)B_0(k)\sigma_1 G(k,\omega-\omega_1)\}\frac{\mathrm{d}\omega}{2\pi} \end{aligned} \tag{4.31}$$

$$\begin{aligned} \chi_{j_1^{new}}^{\prime(1)}(-\omega_1, \omega_1) = -2\left(\frac{ea}{\hbar}\right)^2 \frac{1}{L}\sum_k \int_{-\infty}^{\infty} \mathrm{Tr}\{&A_1^{new}(k)\sigma_3 G(k,\omega-\omega_1) \\ &+ B_1^{new}(k)\sigma_1 G(k,\omega-\omega_1)\}\frac{\mathrm{d}\omega}{2\pi}. \end{aligned} \tag{4.32}$$

Equation (4.31) gives exactly the same result as equation (4.18), which is computed through $J_0J_0$-correlation, while the contribution from $J_1^{new}$ in equation (4.32) can be obtained as

$$\chi_{j_1^{new}}^{\prime(1)}(\omega) = 2\mathrm{i}\frac{e^2(2t_0a)}{\pi\hbar^2}\int_1^{1/\delta} \frac{(1-\eta\delta x^2)^2\,\mathrm{d}x}{[(1-\delta^2x^2)(x^2-1)]^{1/2}x^2}. \tag{4.33}$$

The above term (4.33) actually is $\langle[j_0, j_0]\rangle(\omega = 0)$, suggested by Batistic and Bishop [19].

Using equation (4.18) and equation (4.33), equation (4.30) leads to exactly the same result as equation (3.13) and equation (3.14), which were computed through the $DD$-correlation.

## 5. Discussion

### *5.1. The condition for equivalence between two gauges*

The above examples based on $J_0J_0$- and $DD$-correlations show the importance of the gauge phase factor $F_g$ in optical response theory. The equivalence of the two gauges ($\boldsymbol{E}\cdot\boldsymbol{r}$ and $\boldsymbol{p}\cdot\boldsymbol{A}$) should not be based on exactly the same sets of wavefunctions; they should differ by the gauge phase factor. This crucial point has never been clearly pointed out previously. The

conventional equivalence between the static current–current ($J_0 J_0$-) correlation and the static dipole–dipole ($DD$-) correlation is not maintained; instead, we should consider the induced field currents (IFCs) which are introduced by the gauge phase factor. For example, to consider $\chi^{(1)}$ in the periodic models, we should apply the phase shift from $\boldsymbol{k}$ to $\kappa$ ($\boldsymbol{\kappa} \equiv \boldsymbol{k} + e\boldsymbol{A}/\hbar$) for the basis wavefunctions (or through the creation and annihilation operators in models) to obtain the new current operator $\hat{J}^{new}$. Thus, the equivalence between current–current and dipole–dipole correlations should be understood as the situation when all IFCs are included. This paper's conclusion that the time-dependent gauge phase factor $F_g$ cannot be ignored in the perturbation scheme is also consistent with the results of Langhoff *et al* [36].

### *5.2. The initial distribution function in two gauges*

In optical response theory, one is much more interested in the population of the states than the phase factor. Recently, femtochemistry experiments [39] have been able to reveal the phase factor's effect from the vibrational modes of nuclei as distinct from the population effect (or distribution function) of states. To some extent, the phase factor's influence on the optical response has begun to attract interest. Although our treatment of the optical susceptibilities is still in the framework of non-vibrational nuclei, the gauge phase factor's effect in the theory is still visible in the above computations.

Another important feature caused by the gauge phase factor $F_g$ is the influence on the initial distribution function $f_n(\boldsymbol{k})$ of the two gauges. If all particles are in the ground state $\psi_g^0$ of the unperturbed Hamiltonian $\hat{H}_0$ when the electric field $\boldsymbol{E}$ is applied at time $t = 0$, under the perturbation scheme we can use the set of unperturbed wavefunctions $\{\psi_n^0\}$ with the initial distribution function

$$\begin{aligned} &f_g(\boldsymbol{k}) = 1 \\ &f_n(\boldsymbol{k}) = 0 \qquad \text{for all other } n \neq g. \end{aligned} \tag{5.1}$$

Previously, we used exactly the same initial distribution function, like equation (5.1), for both gauges.

As we showed in the previous sections, the exact wavefunctions for the two gauges should differ by a time-dependent gauge phase factor. This conclusion also holds under the perturbation scheme. Thus, the initial distribution functions in two different gauges are not necessarily exactly the same as equation (5.1), and they should be carefully considered when choosing the basis sets.

Specifically, for the $\boldsymbol{E}\cdot\boldsymbol{r}$ gauge, we should choose unperturbed wavefunctions as the basis to avoid the ZFD directly†; then the initial distribution function can be equation (5.1), since the initial ground state is $\psi_g^0$. But for the $\boldsymbol{p}\cdot\boldsymbol{A}$ gauge, since the new ground state should be $\mathrm{e}^{\mathrm{i}F_g}\psi_g^0$ according to equation (2.3), there are two ways to set the initial distribution function:

(i) If we have already considered the gauge phase factor in our new basis set—that is, if we use $\{\mathrm{e}^{\mathrm{i}F_g}\psi_n^0\}$ as our basis—the distribution function is still like equation (5.1).
(ii) If we still use the unperturbed wavefunction $\{\psi_n^0\}$ as our basis, we should project the initial wavefunction $\mathrm{e}^{\mathrm{i}F_g}\psi_g^0$ onto the basis set $\{\psi_n^0\}$ instead of directly applying equation (5.1) as our initial distribution function.

The latter is much more complicated, since the initial set of distribution functions will be time dependent. In the previous example that we gave, we used the first method under the local

† In fact, this point has not been pointed out clearly previously, even in the Genkins–Mednis approach; $\chi^{(1)}$ is defined in terms of $\boldsymbol{\kappa}$ instead of $\boldsymbol{k}$. Since the basis for the $\boldsymbol{E}\cdot\boldsymbol{r}$ gauge is chosen to be the exact eigenstates under the $\boldsymbol{p}\cdot\boldsymbol{A}$ gauges, this will cause a ZFD directly under $\boldsymbol{E}\cdot\boldsymbol{r}$ gauges even for linear conductivity! Please see [40].

phase approximation†. Both ways of treating the initial distribution function for the $\boldsymbol{p}\cdot\boldsymbol{A}$ gauge are complicated; in this respect, the gauge $\boldsymbol{E}\cdot\boldsymbol{r}$ is much better. This conclusion has actually already been tested in practical applications [2–8, 10, 11, 26–33].

### *5.3. Some previous puzzles*

In explaining the non-equivalence of results for the two gauges, some of the literature has been emphasized by others‡. However, such explanations have lacked direct proof, and the effect of the gauge phase factor has been ignored. The above two subsections provide qualitative explanations for the non-equivalence puzzles for the two gauges that arise in the numerical computations [8, 11, 33]. On the basis of the same set of unperturbed wavefunctions [8], the computations of Bassani *et al* showed that for the $\boldsymbol{E}\cdot\boldsymbol{r}$ gauge the convergence is much faster than for the $\boldsymbol{p}\cdot\boldsymbol{A}$ gauge, and reveal that the distribution equation (5.1) is appropriate for the $\boldsymbol{E}\cdot\boldsymbol{r}$ case but not for the $\boldsymbol{p}\cdot\boldsymbol{A}$ case, which is certainly reasonable in view of our discussion in section 5.2. The $\chi^{(2)}$-computation is a special case [11]§, since some terms could be zero if symmetry is applied [35]. Unfortunately, non-equivalences of the results for the $J_0J_0$-correlation and $DD$-correlation are magnified in the $\chi^{(3)}$-computations. As an example, the spectra of $\chi^{(3)}$ for polyacetylene are different for the two gauges [1, 33]. These differences can be understood qualitatively through considering the gauge phase factor.

## 6. Conclusions

From the computations and discussion above, we concluded that equation (1.2) based on the static current is improper and leads to a ZFD problem; thus, IFCs generated from the gauge phase factor should be included to resolve this difficulty. Generally speaking, the Hamiltonians under the two gauges are not necessarily equivalent unless the gauge phase factor is considered properly through the wavefunctions. Because of the problems regarding the choice of the initial distribution function and the basis sets under the two gauges (section 5.2), it would be very tedious to carry out perturbation computations based on the $\boldsymbol{p}\cdot\boldsymbol{A}$ gauge rather than the $\boldsymbol{E}\cdot\boldsymbol{r}$ gauge. If a careful computation was made based on the concept of the gauge phase factor, the two gauges would lead to equivalent results. Although our computations are chiefly based on a 1D periodic model, it can easily be seen that the chief conclusions of this paper can be expanded to 2D, 3D, and other systems on the basis of the general illustrations in section 2.

## Acknowledgments

One of the authors (M Xu) would like to thank Professor Y Takahashi for intuitive discussions on some concepts in field theory during his visit to Fudan University in 1992. Very helpful discussions with Professor J L Birman, Professor H L Cui, Professor Y R Shen, Professor Z G Soos, Professor C Q Wu, and Dr Z G Yu are acknowledged. This work was supported by the

† Strictly speaking, this gauge phase factor in this paper is an approximate of the diagonal elements in the matrix of the dipole operator $\hat{D}$ [20, 23]. As in the example we give above, the diagonal term is enough for obtaining convergent results in the linear case. But for non-linear cases like $\chi^{(2)}$ and $\chi^{(3)}$, the non-diagonal terms of the $\hat{D}$-matrix will be important for cancelling the ZFD problem through the $JJ$-correlation. Certainly the procedures for including the full gauge phase factor will be complicated and thus impractical for the $JJ$-correlation.

‡ Cohen-Tannoudji *et al* argued that the transition amplitudes at resonance are only correct if exact wavefunctions are used—otherwise different results can be obtained from the two points of view. Please see p 324 of the book [35].

§ In Khurgin and Volsin's work [16], they pointed out the different diagonal matrix elements in the $\chi^{(2)}$-formula that occur for the $\boldsymbol{p}\cdot\boldsymbol{A}$ and $\boldsymbol{E}\cdot\boldsymbol{r}$ gauges.

Chemistry Department, New York University, Project 863 and the National Natural Science Foundation of China (59790050, 19874014).

## References

[1] Wu Weikang 1988 *Phys. Rev. Lett.* **61** 1119
[2] Mahan G D 1990 *Many-Particle Physics* (New York: Plenum) ch 3
[3] Butcher P N and Cotter D 1990 *The Elements of Nonlinear Optics* (Cambridge: Cambridge University Press)
[4] Mahan G D and Subbaswamy K R 1990 *Local Density Theory of Polarizability* (New York: Plenum)
[5] Bloembergen N 1977 *Nonlinear Optics* (New York: Benjamin)
[6] Shen Y R 1984 *The Principles of Nonlinear Optics* (New York: Wiley)
[7] Mukamel S 1995 *Principles of Nonlinear Optical Spectroscopy* (New York: Oxford University Press) and references therein
[8] Bassani F, Forney J J and Quattropani A 1977 *Phys. Rev. Lett.* **39** 1070
[9] Martin P and Schwinger J 1959 *Phys. Rev.* **115** 1342
[10] Haug H and Jauho A P 1996 *Quantum Kinetics in Transport and Optics of Semiconductors* (Berlin: Springer) ch 9.2
[11] Aspnes D E 1972 *Phys. Rev.* B **6** 4648
[12] Rammer J and Smith H 1986 *Rev. Mod. Phys.* **58** 323
[13] Ferry D K and Jacoboni C (ed) 1992 *Quantum Transport in Semiconductors* (New York: Plenum) p 57
[14] Callaway J 1991 *Quantum Theory of the Solid State* 2nd edn (New York: Academic) p 483
[15] Peeters F M, Matulis A, Helm M, Fromherz T and Hilbert W 1993 *Phys. Rev.* B **48** 12 008
[16] Khurgin J B and Volsin P 1998 *Phys. Rev. Lett.* **81** 3777
[17] Maki K and Nakahawa M 1981 *Phys. Rev.* B **23** 5005
[18] Kivelson S, Lee T K, Lin-Liu Y R, Peschel I and Yu L 1982 *Phys. Rev.* B **25** 4173
[19] Batistic I and Bishop A R 1992 *Phys. Rev.* B **45** 5282
[20] Gebhard F, Bott K, Scheidler M, Thomas P and Koch S W 1997 *Phil. Mag.* B **75** 1
[21] Su W P, Schrieffer J R and Heeger A J 1979 *Phys. Rev. Lett.* **42** 1698
Su W P, Schrieffer J R and Heeger A J 1980 *Phys. Rev.* B **22** 2099
[22] Takayama H, Lin-Liu Y R and Maki K 1980 *Phys. Rev.* B **21** 2388
[23] Fradkin E 1991 *Field Theories of Condensed Matter Systems* (New York: Addison-Wesley) p 9
[24] Fann W S, Benson S, Madey J M J, Etemad S, Baker G L and Kajzar F 1989 *Phys. Rev. Lett.* **62** 1492
[25] Heeger A J, Moses D and Sinclair M 1987 *Synth. Met.* **17** 343
[26] Yu J, Friedman B, Baldwin P R and Su W P 1989 *Phys. Rev.* B **39** 12 814
[27] Yu J and Su W P 1991 *Phys. Rev.* B **44** 13 315
[28] Wu C Q and Sun X 1990 *Phys. Rev.* B **41** 12 845
[29] Wu C Q and Sun X 1990 *Phys. Rev.* B **42** R9736
[30] Shuai Z and Brédas J L 1991 *Phys. Rev.* B **44** R5962
[31] Guo F, Guo D and Mazumdar S 1994 *Phys. Rev.* B **49** 10 102
[32] Spano F C and Soos Z G 1993 *J. Chem. Phys.* **99** 9265
[33] Xu Minzhong and Sun Xin 1999 to be submitted
Xu Minzhong and Sun Xin 1999 *Phys. Lett.* A **257** 215
Xu Minzhong and Sun Xin 1999 *Phys. Lett.* A **259** 502
[34] Schiff L I 1955 *Quantum Mechanics* (New York: McGraw-Hill)
[35] Cohen-Tannoudji C, Dupont-Roc J and Grynberg G 1989 *Photons and Atoms* (New York: Wiley)
[36] Langhoff P W, Epstein S T and Karplus M 1972 *Rev. Mod. Phys.* **44** 602
[37] Heeger A J, Kivelson S, Schrieffer J R and Su W P 1988 *Rev. Mod. Phys.* **60** 781 and references therein
[38] Kirchner S, Evertz H G and Hanke W 1999 *Phys. Rev.* B **59** 1825
[39] Zewail A H 1996 *Femtochemistry* ed M Chergui (Singapore: World Scientific)
[40] Genkins V M and Mednis P 1968 *Sov. Phys.–JETP* **27** 609